

An information and field theoretic approach to the grand canonical ensemble

Diego Granziol* and Stephen Roberts†
*Machine Learning Research Group & Oxford-Man Institute,
 University of Oxford, U.K.*

(Dated: March 30, 2017)

We present a novel derivation of the constraints required to obtain the underlying principles of statistical mechanics using a maximum entropy framework. We derive the mean value constraints by use of the central limit theorem and the scaling properties of Lagrange multipliers. We then arrive at the same result using a quantum free field theory and the Ward identities. The work provides a principled footing for maximum entropy methods in statistical physics, adding the body of work aligned to Jaynes’s vision of statistical mechanics as a form of inference rather than a physical theory dependent on ergodicity, metric transitivity and equal a priori probabilities [1]. We show that statistical independence, in the macroscopic limit, is the unifying concept that leads to all these derivations.

PACS numbers: 89.70.Cf, 05.70.-a, 05.20.Gg, 05.30.Ch

I. INTRODUCTION

A. What is Entropy?

The method of maximum entropy (MaxEnt) [2] is a method for generating the least biased estimate possible on the given information, it is maximally non-committal with regard to missing information [3]. Mathematically we maximise the functional

$$S = \int p(\vec{x}) \log p(\vec{x}) d\vec{x} - \sum_i \lambda_i \left[\int p(\vec{x}) f_i(\vec{x}) d\vec{x} - \mu_i \right] \quad (1)$$

with respect to $p(\vec{x})$, where $\langle f_i(\vec{x}) \rangle = \mu_i$ are the constraints. It can be proved under the axioms of consistency, uniqueness and independence, that for constraints in the form of expected values, it is the unique functional who’s extremization represents a valid chain of inference [4]. It has been widely and successfully applied, from condensed matter physics [5] to finance [6, 7]. Along with its path equivalent, maximum caliber (MaxCal) [8], it has been used to successfully derive statistical mechanics, non-relativistic quantum mechanics, Newton’s laws and Bayes’ rule [8, 9].

B. Where do we stand?

In his influential 1957 paper, introducing the concept of MaxEnt, E. T. Jaynes concludes “that statistical mechanics need not be regarded as a physical theory dependent for its validity on the truth of additional assumptions, such as ergodicity, metric transitivity and equal a priori probabilities” [3]. These are considered

irrelevant artefacts of a “dogmatic, single-minded insistence of the frequency theory of probability” [1, 10]. By “know[ing] only the average $\langle E \rangle$ ” or “given the expectation values $\langle E \rangle, \langle n_1 \rangle, \langle n_2 \rangle$ ” Jaynes derives the canonical/quantum grand-canonical ensemble respectively. Yet these crucial constraints, on which the crux of the entire theory is based are seemingly introduced unsuspectingly. We find an argument based on mean square deviations being within reasonable experimental error [1] to neglect higher order moment information. In [11], Lee and Pressé look to derive the uniform micro-canonical ensemble as a consequence of maximizing the Gibbs-Shannon entropy. However, in their derivations we note that the uniform distribution, which they use to describe the microcanonical ensemble, is not a valid density as

$$\sum_i p_i = \sum_i \frac{\delta(E - E_i)}{\Omega(E)} = \frac{1}{\Omega(E)} \neq 1, \quad (2)$$

where we note that the sum over the discrete delta is unity. However, extending the domain of the probability density over both i and E_j we have¹

$$\sum_{i, E_j} p_{i, E_j} = \sum_{i, E_j} \frac{\delta(E_j - E_i)}{\Omega(E)} = \frac{\sum_{E_j} 1}{\Omega(E)} = 1, \quad (3)$$

where $\sum_{E_j} = \Omega(E)$. It is clear to see that the condition imposed, $\sum_i p_i (1 - \delta_{E_i E}) = 0$, leads to this “derived distribution” independent of the method of maximum entropy. This condition is usually justified through the use of ergodicity and metric transitivity [12] (appendix A). We write their p_i as p_{i, E_j} to denote the different states

* diego@robots.ox.ac.uk

† sjrob@robots.ox.ac.uk

¹ We note that it is possible that, instead of the delta distribution, they meant the indicator function $\mathbb{1}(E - E_i)$, this is equivalent to equation (3).

of equal energy.

$$\begin{aligned} \sum_{i,E_j} p_{i,E_j} (1 - \delta_{E_i E_j}) &= 0 \rightarrow 1 = \sum_{i,E_j} p_{i,E_j} \delta_{E_i E_j} \\ &= \sum_{E_j} p_{E_i, i} \rightarrow p_{E_i, i} = \frac{1}{\Omega(E_i)}, p_{E_j \neq i, i} = 0. \end{aligned} \quad (4)$$

We have thus shown, by assuming that p_i is a density, that the constraint implies the Dirac delta without invoking the principles of a Maximum Entropy method. This can be readily seen by using Lagrange multipliers. For a function $f(x) = 0$ with constraints $g_k(x) = 0$ we solve the equations,

$$\begin{aligned} \frac{\partial f(x)}{\partial x_i} - \sum_k \lambda_k \frac{\partial g_k(x)}{\partial x_i} &= 0 \\ \lambda_l \left(\frac{1}{\lambda_l} \frac{\partial f(x)}{\partial x_i} - \frac{\partial g_l(x)}{\partial x_i} - \sum_{k \neq l} \frac{\lambda_k}{\lambda_l} \frac{\partial g_k(x)}{\partial x_i} \right) &= 0. \quad (5) \\ \text{As } \frac{\lambda_l \rightarrow \infty}{\lim} \text{ so } \frac{\partial g_l(x)}{\partial x_i} &= 0. \end{aligned}$$

Hence in the limit of an infinite Lagrange multiplier, the functional form of the entropy is irrelevant. We are simply solving the constraint equation up to a constant addition, as per [11, 13].

C. Application to the Canonical Ensemble

In [11], in order to link the inverse temperature and mean energy constraint, a two step maximisation process is employed, using the decomposition of probabilities and the approximation (due to a large bath size):

$$\log \Omega(E_b - E_i) \approx \log \Omega(E_b) - \beta E_i. \quad (6)$$

This assumption, which is a consequence of the factorisation of micro states (and thus independence), fully constrains the probability density to be of the exponential form.

$$p_i = \frac{\Omega(E_b - E_i)}{\Omega(E_b)} \quad (7)$$

Equating (6) and (7) we obtain,

$$\log p_i = \log(\Omega(E_b - E_i)) - \log(\Omega(E_b)) = -\beta E_i. \quad (8)$$

Generically when considering a large heat bath in contact with an open system, the additive microstates are expanded to first order:

$$\begin{aligned} \Omega(E - E_\nu) &= \Omega(E) \Omega(1 - \frac{E_\nu}{E}) \\ &\propto \Omega(1) + \frac{d\Omega}{dE} \left(\frac{-E_\nu}{E} \right) + \left[\frac{d^2\Omega}{dE^2} \left(\frac{-E_\nu}{E} \right)^2 \dots \right]. \end{aligned} \quad (9)$$

Arguments for this truncation [12] often appeal to the independence of random variables for which

$$\frac{\sqrt{(E - \langle E \rangle)^2}}{E} = \frac{\sqrt{n \text{Var}(\epsilon)}}{n\epsilon} \simeq O\left(\frac{1}{\sqrt{n}}\right) \rightarrow_{n \rightarrow \infty} 0. \quad (10)$$

It is worth noting that it can be readily shown [14] that independence *alone* is enough to constrain the microstate expansion to have the exact form (6). By Liouville's theorem the density (which is time invariant), must be an equation of the integrals of motion (energy and the three components of (angular) momentum). By independence of sub-systems it must be an additive combination of the integrals of motion. By considering the rest frame of a non-rotating system we recover (6) as:

$$\log p_i = \alpha_i + \beta E_i + \gamma \cdot P_i + \delta \cdot \mu_i \rightarrow \alpha_i + \beta E_i. \quad (11)$$

Where $\alpha, \beta, \gamma, \delta$ are constants and E_i, P_i, μ_i are the energy and components of momentum and angular momentum respectively. Equating equations (7) and (11) reproduces equation (6). We note that a corresponding derivation holds for quantum statistics. The quantum analogue of Liouville's theorem is that the statistical matrix (QM equivalent of the distribution function) must commute with the Hamiltonian, in the energy representation it is thus diagonal. Applying this to quasi-closed subsystems one has [14]

$$\log w_n^i = \alpha^i + \beta E_n^i \quad (12)$$

for each subsystem i . This is equivalent to Equation (11).

D. Contributions of this paper

We look to fill a gap in the literature by analytically deriving the mean value constraints (which by the application of MaxEnt lead to the canonical ensembles [1]). We do this via two equivalent methods. We initially proceed by demonstrating that independence follows naturally from the MaxEnt framework and then applying independence to an arbitrary probability distribution over the macro-variables of interest (energy, particle number). By the central limit theorem in the thermodynamic limit ($N \rightarrow \infty$), we recover the mean value constraints. We also derive the same mean value constraints from the underlying quantum field theory. By positing a time-invariant underlying Lagrangian and using the quantum version of Noether's theorem (namely the Ward Identities), one recovers the mean energy constraint (leading to the canonical ensemble by MaxEnt). By positing a free field scalar Lagrangian (no interactions and hence independent) by Ehrenfest's theorem we then recover a particle number constraint (leading to the grand canonical ensemble by MaxEnt). We complete the gap we perceive in the literature, allowing the derivation of the ensembles without the need for additional assumptions such as ergodicity, metric-transitivity and equal a priori probabilities. Further, noting that quantum mechanics has

been shown to be the Bayesian theory extended to complex Hermitian spaces [15] and that nanoscale systems are routinely investigated [16], we regard it as essential to have a fully Bayesian formulation of statistical mechanics, free from infinite mental assembly arguments. By creating an equivalence between free field Lagrangians and statistical independence, we allow the analysis of dependent and small-scale systems and make a tentative link towards the work of Tsallis and its success in optical lattices [17, 18], as well as opening up possible application domains such as finance [19].

II. INDEPENDENCE AND ENTROPY

A. Independence

It is often assumed, due to weak coupling, that the energy (of a statistical mechanical system) is additive and the number of states can thus be factorised. This necessarily implies statistical independence. It can be shown that due to the decomposition of probabilities [20] that the joint entropy is maximal for independent variables;

$$H(X, Y) = H(X) + H(Y) - I(X; Y) \leq H(X) + H(Y), \quad (13)$$

which follows from the fact that the mutual information, $I(X; Y)$, in equation (13), as a KL-divergence (relative entropy), is never negative:

$$\begin{aligned} I(X; Y) &= - \sum_{x, y} p(x, y) \log \frac{p(x)p(y)}{p(x, y)} \\ &= -\mathbb{E} \left[\log \frac{p(x)p(y)}{p(x, y)} \right] \geq -\log \left(\mathbb{E} \left[\frac{p(x)p(y)}{p(x, y)} \right] \right) = 0, \end{aligned} \quad (14)$$

in which $\mathbb{E}(\cdot)$ is the expectation operator. Hence we show that the assumption of independent random variables is consistent with the principle of maximum entropy.

B. Application to Statistical Mechanics

We now consider the application of independence, maximum entropy and the central limit theorem (appendix F) to an ensemble of particles. For a system of N particles, the combinatorial most likely (maximum entropy) state is for them to be independent (this follows from Equation (13)).

Let us assume that there exists some (probability) distribution $p(\vec{x}_1 \dots \vec{x}_n)$, (where \vec{x}_i is the state vector of each individual particle), that describes the complete microstate of the whole ensemble of particles. The entropy

is hence,

$$\begin{aligned} S &= - \int p(\vec{x}_1 \dots \vec{x}_n) \ln p(\vec{x}_1 \dots \vec{x}_n) d(\vec{x}_1 \dots \vec{x}_n) \\ &\quad - \alpha \left[\int p(\vec{x}_1 \dots \vec{x}_n) d(\vec{x}_1 \dots \vec{x}_n) - 1 \right] \\ &\quad - \sum_{i, m} \lambda_{i, m} \left[\int p(\vec{x}_1 \dots \vec{x}_n) \epsilon(\vec{x}_1 \dots \vec{x}_n)^m d(\vec{x}_1 \dots \vec{x}_n) - \mu_{i, m} \right]. \end{aligned} \quad (15)$$

Where S in the above is the system entropy, α and λ are the Lagrange multipliers and $\epsilon(\vec{x}_i)$ is the energy of the system, which depends on all the particles and we characterise by its moments, which are defined by the power m in the above. By being maximally entropic, we consider our particles to be independent and hence the probabilities factorise, thus $p(\vec{x}_1 \dots \vec{x}_n) = \prod_i p(\vec{x}_i)$. We can therefore write $\epsilon(\vec{x}_1 \dots \vec{x}_n) = \sum_i \epsilon(\vec{x}_i)$ and note that all cross correlations are 0, as $\langle \epsilon(\vec{x}_i) \epsilon(\vec{x}_j) \rangle_{i \neq j} = 0$. Further, as there is no way to distinguish one particle from another, they are also identical probability distributions,

$$\begin{aligned} S &= \sum_i \left(- \int p(\vec{x}_i) \ln p(\vec{x}_i) d\vec{x}_i \right. \\ &\quad \left. - \alpha \left[\int p(\vec{x}_i) d\vec{x}_i - 1 \right] \right. \\ &\quad \left. - \sum_{m>1} \lambda_m \left[\int p(\vec{x}_i) \epsilon(\vec{x}_i)^m d\vec{x}_i - \mu_i \right] \right). \end{aligned} \quad (16)$$

Writing the MaxEnt constraints in terms of cumulants and noting that $n = \sum_i$ and $S_N = \sum_i \epsilon_i/n$, we obtain:

$$\sum_i C_m(\vec{\epsilon}_i) = n C_m(\sum_i \vec{\epsilon}_i/n) = \sum_i C_m(S_N), \quad (17)$$

Hence we can write S as

$$\begin{aligned} S &= \sum_i \left(- \int p(\vec{x}_i) \ln p(\vec{x}_i) d\vec{x}_i - \alpha \left[\int p(\vec{x}_i) d\vec{x}_i - 1 \right] \right. \\ &\quad \left. - \sum_{m>1} \lambda_m \left[C_m(S_N) - \mu_i \right] \right) \end{aligned} \quad (18)$$

We now invoke the central limit theorem (appendix F). It can be shown that for N identical independent variables the cumulant of their average varies as $C_m(S_N) = C_m[\epsilon(\vec{x}_i)] N^{1-m}$ and hence for each particle we may solve an effective mean field given as:

$$\begin{aligned} S &= - \int p(\vec{x}) \ln p(\vec{x}) d\vec{x} - \alpha \left[\int p(\vec{x}) d\vec{x} - 1 \right] \\ &\quad - \sum_{m>1} \lambda_m N^{1-m} \left[C_m[\epsilon(\vec{x}_i)] - \mu'_i \right] \end{aligned} \quad (19)$$

So long as the cumulants (which can easily be written in terms of moments, see appendix E) are not infinite, our only non-zero constraint (as μ'_i is a constant) in the

$N \rightarrow \infty$ limit is for $m = 1$. We note that the $m = 0$ term is our normalisation constraint, treated separately with its Lagrange multiplier α . We hence recover a system with a mean value constraint:

$$S = - \int p(\vec{x}) \ln p(\vec{x}) d\vec{x}_i - \alpha \left[\int p(\vec{x}) d\vec{x}_i - 1 \right] - \lambda_1 \left[C_1[\epsilon(\vec{x}_i)] - \mu'_i \right], \quad (20)$$

in which $C_1[\epsilon(\vec{x})] = \langle \epsilon(\vec{x}) \rangle = \int p(\vec{x}) \epsilon(\vec{x}) dx$. We note that through this formulation we have derived the mean energy value constraint via assuming independence (which is maximally entropic) and the central limit theorem (F) in the thermodynamic limit ($N \rightarrow \infty$). Whilst we have applied the method to energy, the formalism naturally extends to other variables of interest, such as particle number. In such extended configurations we have an extra set of Lagrange multipliers in equation (15) over the set of moment expectations $\langle N(x_1 \dots x_n)^m \rangle$. The principle of independence can be readily applied again, in this case to independence between energy and particle number. From this, e.g., we recover a mean particle number constraint along with a mean energy constraint.

The importance of the approach lies in the elegant equivalence between the ensemble and maximum entropy derivations of statistical mechanics under the independence assumption (present in the factorisation of microstates and the additivity of energy).

We further note that convergence can be extended to m -dependent variables [21], where we have a bounded error to the central limit theorem of $\mathcal{O}(N^{-1/2})$ under the condition of a finite 8th moment ($E(X^8) < \infty$) and a finite variance of the sum of the variables. For exponentially decaying variables [22] with a finite 4th moment ($E(X^4) < \infty$), strong mixing and regularity leads to an error to the central limit theorem of $\mathcal{O}(N^{-1/2} \log(N^{1/2}))$. We note that the framework developed in the preceding section is robust to weak dependence and remains exact in the thermodynamic limit ($N \rightarrow \infty$), which formalises the meaning behind sub-systems being “approximately independent” [14].

We also suggest that this formalism is indicative of how to handle small scale systems (where the thermodynamic limit is not appropriate) and highly dependent interacting systems (where the independence assumption no longer holds). Note that in Equation (19), larger higher-order moment information (coming via the cumulant) is equivalent to smaller moment information for a smaller particle number N . This suggests an equivalence between large dependent systems and small independent systems. We leave the rigorous formulation and applications thereof for future work, but note that in many systems of interest the number of observations or agents is far smaller than those of a typical thermodynamic system (namely of order 10^{27}). We further note that for small systems, the total probability mass in the immediate vicinity of the maximum entropy prediction shrinks ($p \propto \exp(NS)$) and so the problem of averaging over

states of lower entropy becomes important. In the discussion we demonstrate how both Tsallis q-exponentials and q-gaussians emerge naturally from classical statistical mechanics in the limits of small heat baths and rest masses.

III. MICROCANONICAL ENSEMBLE

For a time invariant Lagrangian, by Noether’s theorem (see Appendix B) we have a fixed energy. This necessarily implies a delta distribution around the system energy value E_i . Constraining the distribution so that its sum (or integral) is unity leads to:

$$\delta \left(- \int_{E=E_i} p(x) \ln p(x) dx - \alpha \left[\int_{E=E_i} p(x) dx - 1 \right] \right) = 0 \\ \rightarrow p(x) = \exp^{-[1+\alpha]} = \frac{1}{\Omega(E_i)} = \frac{\delta(E - E_i)}{\Omega(E)}. \quad (21)$$

It can be thence shown that for the infinite set of constraints which a distribution of fixed energy must satisfy (all central moments being 0), that the unique distribution to satisfy these constraints is the Dirac delta (see Appendix E). This derivation does not postulate ergodicity, metric transitivity and equal a priori probabilities.

IV. THE CANONICAL ENSEMBLE AS A MAXIMUM ENTROPY DISTRIBUTION

For a system with a conserved mean energy, we maximize the entropy functional with an additional Lagrange multiplier,

$$\delta S = \frac{\partial}{\partial p_j} \left[- \sum_i p_i \ln p_i - \alpha \left(\sum_i p_i - 1 \right) - \beta \left(\sum_i p_i \epsilon_i - \langle \epsilon \rangle \right) \right] = 0 \quad (22)$$

This can be solved to give the canonical ensemble as below,

$$p_i = \frac{e^{-\beta \epsilon_i}}{\sum_i e^{-\beta \epsilon_i}} = \frac{e^{-\beta \epsilon_i}}{Z(\beta)}. \quad (23)$$

Where the Lagrange multiplier β is $\langle \epsilon \rangle^{-1}$ from the mean value constraint (the third expression in Equation (22)).

A. Mean energy constraint

In Section II B we derived the mean energy constraint explicitly using arguments of independence and the central limit theorem in the thermodynamic limit. We now show that equivalently, the constraint can be derived from the underlying quantum field theory. Generalising Noether’s theorem from classical field theory to give

a quantum field theoretic version, we derive the Ward identities (Appendix C). The end result is that we no longer have a conserved current but a conserved *expected* current,

$$\partial^\mu \langle j_\mu(x) \rangle = 0. \quad (24)$$

In the case of continuous time translation symmetry we would now also have a conserved energy expectation. Under the MaxEnt framework this also leads to eqn (23).

V. GRAND CANONICAL ENSEMBLES FROM MAXIMUM ENTROPY

Generalising Equation (22) to a two-dimensional probability distribution over both energy and particle number, with mean constraints on both energy and particle number we have,

$$\begin{aligned} \frac{\partial}{\partial p_{k,l}} \left[- \sum_{i,j} p_{i,j} \ln(p_{i,j}) - \alpha \left(\sum_{i,j} p_{i,j} - 1 \right) \right. \\ \left. - \beta \left(\sum_{i,j} p_{i,j} \epsilon_i - \langle \epsilon \rangle \right) - \gamma \left(\sum_{i,j} p_{i,j} n_j - \langle n \rangle \right) \right] = 0. \end{aligned} \quad (25)$$

therefore,

$$p_{ij} = \frac{e^{-\beta(\epsilon_i + \frac{\gamma}{\beta} n_j)}}{\sum_{i,j} e^{-\beta(\epsilon_i + \frac{\gamma}{\beta} n_j)}}. \quad (26)$$

upon making the substitution $\gamma/\beta = -\mu$ we arrive at the celebrated grand canonical ensemble probability distribution,

$$p_{ij} = \frac{e^{-\beta(\epsilon_i - \mu n_j)}}{Z(\beta)}. \quad (27)$$

A. Conserved mean particle number

In this section, we derive the constraint for a fixed particle number using Ehrenfest's theorem and a free field theory (which posits underlying independence between the particles). We retain the constant energy expectation from the previous section.

B. Ehrenfest's Theorem

We start by quoting Ehrenfest's theorem,

$$\begin{aligned} \frac{d}{dt} \langle \hat{O} \rangle &= \frac{d}{dt} \int \psi^* \hat{O} \psi \\ &= \int \psi \frac{\partial \hat{O}}{\partial t} \psi + \int \frac{\partial \psi^*}{\partial t} \hat{O} \psi + \int \psi^* \hat{O} \frac{\partial \psi}{\partial t}. \end{aligned} \quad (28)$$

Using Schroedinger's equation, $H|\psi\rangle = i\hbar \frac{\partial |\psi\rangle}{\partial t}$, we obtain,

$$\frac{d}{dt} \langle \hat{O} \rangle = \left\langle \frac{\partial \hat{O}}{\partial t} \right\rangle + \frac{1}{i\hbar} \langle \psi | [\hat{O}, \hat{H}] | \psi \rangle. \quad (29)$$

C. Scalar Klein-Gordon

The simplest relativistic-free theory is that defined by the classical Klein-Gordon (KG) equation, in which the Lagrangian, \mathcal{L} , is given as,

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \phi)^2 - \frac{1}{2} m^2 \phi^2. \quad (30)$$

It can be shown by writing out the fields, number operator and conjugate momenta in terms of the creation and annihilation operators (see Appendix C 3) that for a free field theory, the Hamiltonian commutes with the number operator, i.e. $[N, H] = 0$. Applying Ehrenfest's theorem to the number operator we see that,

$$\frac{d}{dt} \langle \hat{N} \rangle = \left\langle \frac{\partial \hat{N}}{\partial t} \right\rangle + \frac{1}{i\hbar} \langle \psi | [\hat{N}, \hat{H}] | \psi \rangle = 0, \quad (31)$$

as the number operator does not explicitly depend on time and its commutator is 0. Hence,

$$\langle N \rangle = k. \quad (32)$$

This completes our derivation. We see that for a time-invariant, non-interacting Lagrangian (one in which the number operator commutes with the Hamiltonian), in the large number limit (this causes the maximum entropy solution to dominate the probability mass), the exponentially most likely solution is the grand canonical ensemble.

VI. EXTENSION TO INTERACTING COMPLEX LAGRANGIANS

A Lagrangian of the form $\mathcal{L} = \partial_\mu \psi^* \partial^\mu \psi - V(|\psi|^2)$ is invariant under the change $\psi \rightarrow e^{i\alpha} \psi$ or $\delta\psi = i\alpha\psi$. $\delta\psi^* = i\alpha\psi^*$. Due to the Ward identities (see Appendix C), expanding the complex field operators and the classical field momentum as a sum of plane waves (see Appendix D) we have a conserved mean quantum charge $\langle Q \rangle = \langle N_c - N_b \rangle$ in the interacting theory. This means, in the thermodynamic limit, following a similar argument to that of Section IV we would expect an exponential over the particle number of the form,

$$P_{b,c} = \frac{1}{Z} \exp[(N_c - N_b)/\langle N_c - N_b \rangle]. \quad (33)$$

And hence we see that even for strongly interacting systems with symmetries we expect to be able to characterise these systems by exponential distributions over the conserved expectations induced by the symmetries.

VII. CONCLUSION AND DISCUSSION

In this paper we have derived all the classical statistical mechanical ensembles without the need for ergodicity, metric transitivity or the postulate of equal a priori probabilities. By further exploiting the independence assumption (already present in the ensemble theory approach) and applying the central limit theorem in the thermodynamic limit, we derived the mean value constraints. By the method of maximum entropy these constraints reproduce the canonical and grand-canonical ensembles. The mean value constraints were also derived by the application of the Ward identities and Ehrenfest's theorem to a time invariant free-field Lagrangian. That we arrive at the same prediction under two completely different formalisms builds on the work of Jaynes [1] and Benavoli [15], displaying the extent to which Physics can be considered as a theory of inference. Open questions still remain, such as how to extend the formalism to predict probability distributions of systems characterised by non-time invariant interacting Lagrangians. Our result in Section VI, exploiting the $U(1)$ symmetry for an interacting Lagrangian, perhaps offers a starting point. Analogies between higher-order Lagrangians and statistical dependence have also not been explored in this paper.

By considering statistical results for weakly dependent variables [21, 22], we have shown the robustness of the conclusion under certain types of dependence in the thermodynamic limit. Unanswered questions include how to cope in situations of extreme dependence and in the small number limit. We further note that the informativeness of the MaxEnt solution is dependent on the sharpness of the predictive distribution around its most entropic value; for small systems the contribution of lower entropy solutions will be significant and should be taken into account.

The value of Tsallis entropies: The BSG entropic functional (Equation (1)) has been extended to a q-entropic functional [23]. It has been used to better model financial returns and price options [18] as well as predict the distribution in optical lattices [17]. Its free parameter, q , defines the fluctuations of the generalised temperature (following a gamma distribution) [24] and can be derived from Langevin equations with fluctuating forces [25]. It has been criticised for failing to reproduce basic thermodynamical laws such as Kirchoff's and Boltzmann's law for anything other than $q = 1$ (the value which reduces it to the BSG entropy) [26] and for introducing unwanted bias not warranted by the data [27]. We simply show that, in a much more restrictive sense ($q \leq 1$), the Tsallis q-exponential and q-Gaussian can be derived from generic statistical mechanics in the limit of small particle number and negligible rest mass.

We start by writing the total number of microstates of

an ideal gas as,

$$\Omega(E, N, V) = \frac{1}{h^{3N}} \int_V \Pi_i d\vec{x}_i \int_E \Pi_i d\vec{p}_i = \frac{V^N}{h^{3N}} \int_E \Pi_i d\vec{p}_i. \quad (34)$$

The momenta are constrained by the usual relation $(1/2m) \sum_i (\vec{p}_i^2) = E$, so we have a $3N$ dimensional hypersphere of radius $\sqrt{2mE}$ and accessible momentum space is RdR . Combining this with the equation of volume and surface area of an n -dimensional hypersphere,

$$V_n(R) = \frac{\pi^{n/2}}{(n/2)!} R^n, \quad S_n(R) = \frac{2\pi^{n/2}}{\gamma(n/2)!} R^{n-1}, \quad (35)$$

we have

$$\Omega(E, N, V) = \frac{V^N}{h^{3N}} \frac{(2\pi m E)^{3N/2}}{(3N/2 - 1)!} \frac{\delta E}{E}. \quad (36)$$

Hence, as discussed in [9],

$$p_i = \frac{\Omega(E_b - E_i)}{E_b} \propto (1 - E_i/E_b)^{3N/2-1}, \quad (37)$$

which is equivalent to the Tsallis q-exponential

$$p_i \propto (1 - \lambda E_i)^{1/1-q}. \quad (38)$$

We note that the constraint $\sum_i p_i^2 = 2mE$ is a special case of the special relativistic invariant equation $E^2 = p^2 c^2 + M_0^2 c^4$, where E is the energy of the particle, p is the momentum, M_0 the particle rest mass and c is the speed of light. Adopting natural units where $c = 1$ we have $\sum_i p_i^2 = E^2 - M_0^2$. Analogously to (37) we generate a distribution which is q-exponential in the Newtonian limit and q-Gaussian² in the ultra relativistic limit.

$$p_i \propto (1 - \lambda(E_i^2 - M_{0,i}^2))^{1/1-q}. \quad (39)$$

The notion of mass emerging generally as a constraint on mean path square displacement (imposing smoothness) [8] gives insight as to which distribution may be appropriate in generic inference. Whilst it is promising that we have made a link between q distributions and the departure from the thermodynamic limit ($N \not\rightarrow \infty$) and the Newtonian limit $((E^2 - M_0^2 c^4)/c^2 \not\approx 2mE)$, we note that the for the above formalism

$$q = 1 - (DN/2 - 1)^{-1} \leq 1, \quad (40)$$

where D is the dimension of the system. Hence none of the q-distributions describing financial, optical lattice or turbulent phenomena [17, 18, 25] can be described by this. However, it derives the probability distribution of a bounded variable, used in Biology, Engineering and Physics [28] as a function of bath size.

² Equation 39 is not the q-Gaussian as $E \not\ll 0$. It is a truncated half q-Gaussian.

Appendix A: Ergodicity and Metric Transitivity

Generically, the functional form of the micro-canonical ensemble is derived under the assumption that for an isolated (constant energy) system in equilibrium, all accessible microstates are equally likely. This restricts us to the probability distribution [12] given by,

$$P_s = \frac{\delta(E_s - U)}{\Omega(E)} = \frac{1}{\Omega(U)}. \quad (\text{A1})$$

This starting point can be derived by assuming ergodicity and metric transitivity. Metric transitivity ensures that the trajectories can move freely on the energy surface (all states of equal energy are accessible), whilst ergodicity ensures that the time the trajectory spends in a region phase space is proportional to its volume (we spend an equal amount of time in all the states of identical energy). As we have nothing to distinguish one region from another, so the best choice we can make is to assume that the probability $P(R_e)$ of finding the system in R_E is equal to the fraction of the energy surface occupied by R_E (this is known as the equal a priori postulate of statistical mechanics). Thus,

$$P(R_E) = \frac{1}{\sum(E)} \int_{R_E} dS_E = \frac{\sum(R_E)}{\sum(E)}, \quad (\text{A2})$$

which when correctly normalised gives (on the accessible energy surface),

$$\rho(X^N, S_E) = \frac{1}{\sum(E)}. \quad (\text{A3})$$

Appendix B: Noether's Theorem

For any continuous symmetry we can work infinitesimally using $\delta\phi_a(x) = X_a(\phi)$. To keep the action invariant we must change the Lagrangian by a total derivative. This is because the equations of motion are derived from extremising the action and hence for fixed boundaries any added constant value $F(b) - F(a)$ is irrelevant. Hence $\partial\mathcal{L} = \partial_\mu F^\mu$, for some set of functions $F^\mu(\phi)$. For an arbitrary transformation of the fields $\partial\phi_a$,

$$\begin{aligned} \partial\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi_a} \delta\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} \partial_\mu(\delta\phi_a) \\ \partial\mathcal{L} &= \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} \right] \delta\phi_a + \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} \delta\phi_a \right). \end{aligned} \quad (\text{B1})$$

The first term in the above is 0 when the equations of motion are satisfied, so by equating this with our total derivative we have,

$$\partial_\mu j^\mu = \partial_\mu \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} X_a(\phi) - F^\mu(\phi) \right] = 0. \quad (\text{B2})$$

1. Energy-Momentum Tensor

If we consider the infinitesimal translation, $x^v \rightarrow x^v - \epsilon^v$, $\phi_a(x) \rightarrow \phi_a(x) + \epsilon^v \partial_v \phi_a(x)$, then we have a variation in the Lagrangian of $\mathcal{L}(x) \rightarrow \mathcal{L}(x) + \epsilon^v \partial_v \mathcal{L}(x) = \mathcal{L}(x) + \epsilon^v \partial_\mu (\delta_v^\mu \mathcal{L}(x))$. Using Noether's theorem we get four conserved currents

$$j_v^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} \partial_v \phi_a - \delta_v^\mu \mathcal{L} \equiv T_v^\mu. \quad (\text{B3})$$

For a Lagrangian of the form for a real scalar field, $\frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2$, we have an energy momentum tensor

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \eta^{\mu\nu}\mathcal{L}. \quad (\text{B4})$$

The 0th component of the current is conserved, giving for conserved momenta

$$P^i = \int d^3x T^{0i}. \quad (\text{B5})$$

For $i = 0$ (the index of time) we have energy and for $i = 1, 2, 3$ we have the momenta in the x, y, z directions.

Appendix C: Mean energy from Quantum Mechanics

1. From Classical to Quantum

We now generalise Noether's seminal result from classical field theory to give a quantum field theoretic version of the result. Completing the same procedure as before, in the notation of Riemann geometry we have:

$$\begin{aligned} \delta_\epsilon S[\phi] &= 0 = - \int_M *j \wedge d\epsilon = \int_M g^{\mu\nu} j_\mu(x) \partial_\nu \epsilon(x) \sqrt{g} d^d x \\ &\rightarrow d * j = \partial_\nu (\sqrt{g} g^{\mu\nu} j_\nu) = 0, \end{aligned} \quad (\text{C1})$$

where $*$ represents the Hodge dual, \wedge is the wedge product and $d\epsilon$ is the exterior derivative.

2. Ward's Identities

Supposing that a local transformation $\phi \rightarrow \phi'(\phi)$ leaves the product of the action and the path integral measure invariant, so:

$$\mathcal{D}\phi e^{-S_\Lambda^{eff}[\phi]} = \mathcal{D}\phi' e^{-S_\Lambda^{eff}[\phi']}. \quad (\text{C2})$$

On a compact manifold M , by relabelling ϕ by ϕ' and using the identity (C2), we have that the correlation functions obey the relations:

$$\langle \mathcal{O}_1(\phi(x_1)) \dots \mathcal{O}_n(\phi(x_n)) \rangle = \langle \mathcal{O}_1(\phi'(x_1)) \dots \mathcal{O}_n(\phi'(x_n)) \rangle. \quad (\text{C3})$$

Hence symmetries impose selection rules on the operators we can insert, if we wish to obtain a non-zero correlator. Computing the path integral in terms of the primed quantities where $\phi' = \phi + \delta_\epsilon \phi$ yields:

$$\begin{aligned} \mathcal{Z} &= \int \mathcal{D}\phi' e^{-S_\Lambda^{eff}[\phi']} = \int \mathcal{D}[\phi + \delta_\epsilon \phi] e^{-S_\Lambda^{eff}[\phi + \delta_\epsilon \phi]} \\ &= \int \mathcal{D}\phi [1 + \delta_\epsilon S] e^{-S[\phi]} (1 - \delta_\epsilon \phi \delta_\epsilon S \dots). \end{aligned} \quad (\text{C4})$$

The last term in (C4) is second order and is thus neglected. Using Equation (C1) this reduces to,

$$\begin{aligned} \int \mathcal{D}\phi' e^{-S_\Lambda^{eff}[\phi']} &= \int \mathcal{D}\phi e^{-S_\Lambda^{eff}[\phi]} \left[1 - \int_M *j \wedge d\epsilon \right] \\ &= \int \mathcal{D}\phi e^{-S_\Lambda^{eff}[\phi]}. \end{aligned} \quad (\text{C5})$$

Hence, as the partition function (cast here in path integral form) calculates sample expectations, we obtain:

$$0 = - \int_M * \langle j(x) \rangle \wedge d\epsilon = \int_M d * \langle j(x) \rangle \rightarrow \partial^\mu \langle j_\mu(x) \rangle = 0. \quad (\text{C6})$$

In the case of continuous time translation symmetry we would now have a conserved energy expectation. It is worth noting that if Equation (C6) applied to the conserved charge of Noether's theorem (see Appendix B) we obtain a mean conserved charge.

3. Scalar Klein Gordon system

The simplest relativistic free theory leads to the classical Klein-Gordon (KG) equation for the Lagrangian:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2 \phi^2, \quad (\text{C7})$$

$$\mathcal{H} = p\dot{q} - \mathcal{L} = (\partial_0 \phi)^2 - \mathcal{L} = \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla \phi)^2 + \frac{1}{2}m^2 \phi^2. \quad (\text{C8})$$

Writing out the fields and their conjugate momenta in terms of the creation and annihilation operators, $w_p = \sqrt{|p|^2 + m^2}$ and $p_\mu = (w_p, \mathbf{p})$, $[a(p), a(q)] = 0$ & $[a(p), a^\dagger(q)] = (2\pi)^3 2w_p \delta^3(\mathbf{p} - \mathbf{q})$ we obtain

$$\phi(x, t) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2w_p} (a(p) e^{-ipx} + a^\dagger(p) e^{ipx}),$$

$$\pi(x, t) = \dot{\phi}(x, t) = -i \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2} (a(p) e^{-ipx} - a^\dagger(p) e^{ipx}).$$

By substituting these relations into the Hamiltonian of Equation (C8) and by placing all the annihilation operators on the left (normal ordering) we obtain the Hamiltonian in terms of the creation/annihilation operators,

$$: H : = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2} a^\dagger(p) a(p), \quad (\text{C9})$$

which commutes with the number operator $[N, H] = 0$. Hence,

$$N = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2w_p} a^\dagger(p) a(p). \quad (\text{C10})$$

This can be seen to count the number of particles in a state by the use of its commutation relation $[N, a^\dagger(q)] = a^\dagger(q)$ on the state $N|p_1, \dots, p_n\rangle$ where $n|0\rangle = 0$.

Appendix D: U(1) symmetry for complex Lagrangians

A Lagrangian of the form $\mathcal{L} = \partial_\mu \psi^* \partial^\mu \psi - V(|\psi|^2)$ is invariant under the change $\psi \rightarrow e^{i\alpha} \psi$ or $\delta\psi = i\alpha\psi$. $\delta\psi^* = i\alpha\psi^*$. Taking our general equation for the conserved current, we obtain an unchanged Lagrangian, hence $F^\mu = 0$. However, the Lagrangian now depends on both ψ and its conjugate ψ^* thus we have the expression for $\partial_\mu j^\mu$ given as,

$$\partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \delta\psi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^*)} \delta\psi^* \right] = \partial_\mu \left[i(\partial^\mu \psi^*) \psi - i\psi^* (\partial^\mu \psi) \right]. \quad (\text{D1})$$

This acts as a conservation of charge or particle number. By expanding the complex field operator and the classical field momentum as a sum of plane waves we obtain,

$$\psi = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\vec{p}}}} \left(b_{\vec{p}} e^{+i\vec{p} \cdot \vec{x}} + c_{\vec{p}}^\dagger e^{-i\vec{p} \cdot \vec{x}} \right),$$

$$\pi = \int \frac{d^3 p}{(2\pi)^3} i \sqrt{\frac{E_{\vec{p}}}{2}} \left(b_{\vec{p}}^\dagger e^{-i\vec{p} \cdot \vec{x}} - c_{\vec{p}} e^{+i\vec{p} \cdot \vec{x}} \right), \quad (\text{D2})$$

which satisfy the relations $[\psi(\vec{x}), \pi(\vec{y})] = i\delta^3(\vec{x} - \vec{y})$ and $[\psi(\vec{x}), \pi^\dagger(\vec{y})] = 0$. This is thus equivalent to the relations $[b_{\vec{p}, \vec{q}}, c_{\vec{p}, \vec{q}}] = (2\pi)^3 \delta^3(\vec{p} - \vec{q})$, the same for c and all other commutators (or anti-commutators if the underlying statistics are Fermi-Dirac instead of Bose-Einstein) being 0. The classical field momentum $\pi = \partial \mathcal{L} / \partial \dot{\psi} = \psi^*$. Hence the conserved classical charge is,

$$Q = i \int d^3 x (\dot{\psi}^* \psi - \psi^* \dot{\psi}) = i \int d^3 x (\pi \psi - \psi^* \pi^*), \quad (\text{D3})$$

which after normal ordering becomes,

$$Q = \int \frac{d^3 p}{(2\pi)^3} (c_{\vec{p}}^\dagger c_{\vec{p}} - b_{\vec{p}}^\dagger b_{\vec{p}}) = N_c - N_b, \quad (\text{D4})$$

where we have defined the *number operator* $N_{a\vec{p}}$ in the n -particle Hilbert space (Fock Space) as,

$$N_{a\vec{p}} = \int \frac{d^3 p}{(2\pi)^3} a_{\vec{p}}^\dagger a_{\vec{p}}. \quad (\text{D5})$$

As a consequence of the Ward identities (appendix C), we have a conserved mean quantum charge $\langle Q \rangle = \langle N_c - N_b \rangle$ in the interacting theory.

Appendix E: Uniqueness of the Delta Distribution

The constraint of a fixed energy necessarily means that there is a fixed mean energy (value E say) and fixed higher order central moments (of value 0). A system with fixed energy classically corresponds to the micro canonical ensemble, this can be derived as follows. Instead of directly maximising the entropic functional,

$$S = - \int p_\epsilon \ln p_\epsilon - \alpha \left[\int p_\epsilon - 1 \right] - \sum_{i \geq 1} \lambda_i \left[\int p_\epsilon (\epsilon - \langle E \rangle)^i = 0 \right], \quad (\text{E1})$$

(in which we have left out the $d\epsilon$ in the functional integrals to unclutter the notation). It can be shown that the delta distribution, satisfies the constraint of normalising to 1, has a mean of E and has all higher order central moments being 0, i.e.

$$\int (E - E')^n \delta(E - E') dE' = (E - E)^n = 0. \quad (\text{E2})$$

To prove that it is the only distribution that satisfies this infinite number of constraints, we use the definition of cumulants, $C_n(E)$, and characteristic functions, $\phi_E(k)$, to obtain,

$$\phi_E(k) = \langle e^{ikE} \rangle \equiv \int P_E(\epsilon) e^{ik\epsilon} d\epsilon = \exp \left(\sum_{n=1} C_n(\epsilon) \frac{(ik)^n}{n!} \right), \quad (\text{E3})$$

where the probability distribution P_E is normalised to 1, $C_1 = \langle E \rangle$ and $C_2 = \langle E^2 \rangle - \langle E \rangle^2$ etc. Setting all of the higher order cumulants other than C_1 to 0 yields,

$$\int P_E(\epsilon) e^{ik\epsilon} d\epsilon = \exp(\langle E \rangle ik). \quad (\text{E4})$$

Differentiating the functional with respect to e^{ikE} gives,

$$\begin{aligned} \frac{\delta \phi_E(k)}{\delta e^{ik'E}} &= P_X(x) \delta(k - k') = \delta(k - k') \delta(\langle E \rangle - E) \\ &\Rightarrow P_X(x) = \delta(\langle E \rangle - E). \end{aligned} \quad (\text{E5})$$

Bell Polynomials: To show that this is equivalent to assuming that all the higher order central moments are 0, we write the n -th cumulant in terms of the central moments,

$$\begin{aligned} C_{n>1} &= \sum_{k=1}^n (-1)^{k-1} (k-1)! B_{n,k}(0, \kappa_2, \kappa_{n+1-k}) \\ &= \sum_{k=1}^n (-1)^{k-1} (k-1)! \sum \frac{n!}{\prod_{a=1}^{n+1-k} j_a!} \left(\frac{\kappa_a}{n!} \right)^{j_a} \end{aligned} \quad (\text{E6})$$

- [1] Edwin T. Jaynes. *Foundations of Probability Theory and Statistical Mechanics*, pages 77–101. Springer Berlin Heidelberg, Berlin, Heidelberg, 1967.
- [2] Steve Pressé, Kingshuk Ghosh, Julian Lee, and Ken A. Dill. Principles of maximum entropy and maximum cal-

$$\forall j_a \sum_{a=1}^{n+1-k} j_a = k, \sum_{a=1}^{n+1-k} a j_a = n.$$

Where we have expanded the Bell polynomials. We note that the cumulant C_n can be written as a combination of a multinomial power series in the central moments (κ_a), which we have specified to be 0 and hence the $C_n = 0$ for $n \geq 2$. This concludes our derivation of the micro-canonical ensemble. Hence, for a time invariant Lagrangian, which by Noether's classical field theoretic result implies a fixed total energy, the only possible solution is the Dirac delta distribution.

Appendix F: Central Limit Theorem

Consider $S_N \equiv Y_N/N$ for large N , the distribution of S_N converges towards a Gaussian with mean $\langle X \rangle$ and standard deviation σ/\sqrt{N} , provided higher order moments are finite. **Proof:** via the characteristic function,

$$\begin{aligned} \phi_{S_N}(k) &= \langle e^{ikS_N} \rangle \\ &= \langle e^{i\frac{k}{N} \sum_{j=1}^N X_j} \rangle = \langle e^{i\frac{k}{N} X} \rangle^N = [\phi_X(k/N)]^N, \end{aligned} \quad (\text{F1})$$

where we have used the properties of independence and identity. In terms of cumulants,

$$\begin{aligned} &\exp \left(\sum_{m=0}^{\infty} \frac{(ik)^m}{m!} C_m(S_N) \right) \\ &= \exp \left(N \sum_{m=0}^{\infty} \frac{(ik/N)^m}{m!} C_m(S_X) \right), \end{aligned} \quad (\text{F2})$$

thus

$$C_M(S_N) = C_m(X) N^{1-m}. \quad (\text{F3})$$

Finite higher order moments all tend to 0 as $N \rightarrow \infty$. As the Gaussian is uniquely defined by its mean and variance this becomes the limit distribution. This can be further seen by taking the inverse Fourier transform of the characteristic function and keeping the first two cumulants,

$$P_{S_N}(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iks} \phi_{S_N}(k) dk. \quad (\text{F4})$$

- iber in statistical physics. *Rev. Mod. Phys.*, 85:1115–1141, Jul 2013.
- [3] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630, May 1957.

- [4] John Shore and Rodney Johnson. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on information theory*, 26(1):26–37, 1980.
- [5] Adom Giffin, Carlo Cafaro, and Sean Alan Ali. Application of the maximum relative entropy method to the physics of ferromagnetic materials. *Physica A: Statistical Mechanics and its Applications*, 455:11 – 26, 2016.
- [6] Cassio Neri and Lorenz Schneider. Maximum entropy distributions inferred from option portfolios on an asset. *Finance and Stochastics*, 16(2):293–318, 2012.
- [7] Peter W Buchen and Michael Kelly. The maximum entropy distribution of an asset inferred from option prices. *Journal of Financial and Quantitative Analysis*, 31(01):143–159, 1996.
- [8] Diego González, Sergio Davis, and Gonzalo Gutiérrez. Newtonian dynamics from the principle of maximum caliber. *Foundations of Physics*, 44(9):923–931, 2014.
- [9] A Caticha. Entropic inference and the foundations of physics (monograph commissioned by the 11th brazilian meeting on Bayesian statistics–ebeb-2012, 2012.
- [10] R. D. Rosenkrantz. *Brandeis Lectures (1963)*, pages 39–76. Springer Netherlands, Dordrecht, 1989.
- [11] Julian Lee and Steve Pressé. Microcanonical origin of the maximum entropy principle for open systems. *Phys. Rev. E*, 86:041126, Oct 2012.
- [12] David Chandler. *Introduction to Modern Statistical Mechanics*. Oxford University Press, 1987.
- [13] Julian Lee. Microcanonical origin of the maximum entropy principle for open systems. *arxiv*, 2012.
- [14] L. D. Landau, Lifshitz E. M., and Pitaevskii L. P. *Statistical physics*. Elsevier, 2012.
- [15] Alessio Benavoli, Alessandro Facchini, and Marco Zafalon. Quantum mechanics: The Bayesian theory generalized to the space of Hermitian matrices. *Physical Review A*, 94(4), Oct 2016.
- [16] Carlos Bustamante, Zev Bryant, and Steven B. Smith. Ten years of tension: single-molecule DNA mechanics. *Nature*, 421(6921):423427, 2003.
- [17] P. Douglas, S. Bergamini, and F. Renzoni. Tunable Tsallis distributions in dissipative optical lattices. *Physical Review Letters*, 96(11), 2006.
- [18] Lisa Borland. Option pricing formulas based on a non-Gaussian stock price model. *Physical Review Letters*, 89(9), Jul 2002.
- [19] S. Picoli Jr., R. S. Mendes, L. C. Malacarne, and R. P. B. Santos. q-distributions in complex systems: a brief review. *Brazilian Journal of Physics*, 39(2a):468474, 2009.
- [20] R. T. Cox. Probability, frequency and reasonable expectation. *American Journal of Physics*, 14(1):1–13, 1946.
- [21] Charles Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*, pages 583–602, Berkeley, Calif., 1972. University of California Press.
- [22] A. N. Tikhomirov. On the convergence rate in the central limit theorem for weakly dependent random variables. *Theory of Probability & Its Applications*, 25(4):790809, 1981.
- [23] Constantino Tsallis. *Introduction to nonextensive statistical mechanics: approaching a complex world*. Springer, 2009.
- [24] G. Wilk and Z. Wodarczyk. Interpretation of the nonextensivity parameter q in some applications of Tsallis statistics and Lévy distributions. *Physical Review Letters*, 84(13):27702773, 2000.
- [25] Christian Beck. Dynamical foundations of nonextensive statistical mechanics. *Physical Review Letters*, 87(18), Oct 2001.
- [26] Michael Nauenberg. Critique of q -entropy for thermal statistics. *Physical Review E*, 67(3), 2003.
- [27] Steve Press, Kingshuk Ghosh, Julian Lee, and Ken A. Dill. Nonadditive entropies yield probability distributions with biases not warranted by the data. *Physical Review Letters*, 111(18), Jan 2013.
- [28] Alberto d’Onofrio. Bounded noises in physics, biology, and engineering. *Modeling and Simulation in Science, Engineering and Technology*, 2013.